

SEMI-AUTOMATIC SEGMENTATION OF AUTOSOMAL DOMINANT POLYCYSTIC KIDNEYS USING RANDOM FORESTS

Kanishka Sharma^{1,2}, Loïc Peter², Christian Rupprecht^{2,3}, Anna Caroli¹, Lichao Wang², Andrea Remuzzi^{1,4}, Maximilian Baust², Nassir Navab^{2,3}

¹ IRCCS - Istituto di Ricerche Farmacologiche “Mario Negri”, Bergamo, Italy

² Computer Aided Medical Procedures, Technische Universität München, Germany

³ Computer Aided Medical Procedures, Johns Hopkins University, USA

⁴ University of Bergamo, Bergamo, Italy

ABSTRACT

This paper presents a method for 3D segmentation of kidneys from patients with autosomal dominant polycystic kidney disease (ADPKD) and severe renal insufficiency, using computed tomography (CT) data. ADPKD severely alters the shape of the kidneys due to non-uniform formation of cysts. As a consequence, fully automatic segmentation of such kidneys is very challenging. We present a segmentation method with minimal user interaction based on a random forest classifier. One of the major novelties of the proposed approach is the usage of geodesic distance volumes as additional source of information. These volumes contain the intensity weighted distance to a manual outline of the respective kidney in only one slice (for each kidney) of the CT volume. We evaluate our method qualitatively and quantitatively on 55 CT acquisitions using ground truth annotations from clinical experts.

Index Terms— Computed Tomography, Image Segmentation, Machine Learning, Random Forests

1. INTRODUCTION

Autosomal dominant polycystic kidney disease (ADPKD) is the most common inherited cystic kidney disease. It is characterized by progressive enlargement of the kidneys caused by sustained development and expansion of renal cysts [1], eventually inducing end stage renal disease (ESRD) in majority of the patients. In addition, ADPKD is associated with a number of extra renal complications such as early hypertension, cerebral aneurysms, heart valve defects, and presence of cysts in other organs like the liver and pancreas. For the purpose of diagnosis, Ultrasonography (US) is usually performed for pre-symptomatic identification of ADPKD, followed by Computed Tomography (CT) or Magnetic Resonance Image (MRI) acquisitions, offering higher resolution [2]. So far,

there are no existing proven treatments for ADPKD, therefore an effective disease-modifying drug would have important implications for patients. Development of efficient computational means for monitoring kidney expansion by reliably quantifying total kidney volumes (TKV) is of crucial importance for assessment of ADPKD progression and evaluating the efficacy of novel therapies. Therefore, fast, accurate and user-friendly segmentation methods are required for TKV computation in ADPKD. However, segmentation of pathological kidneys is not straightforward due to non-uniform renal cyst growth in ADPKD, leading to high variability in terms of organ shape and size. In addition to this, adjacent liver cysts, typically exhibiting similar intensity statistics as kidney cysts [4], further hamper accurate kidney delineation. In case of late stage ADPKD patients, such as those with severe renal insufficiency, hemorrhagic kidney cysts lead to extra complication of high intensity variability with respect to regular fluid-filled cysts. A fully automated segmentation approach allowing reliable and quick TKV estimation remains a challenging task. In this paper, we present a semi-automatic segmentation method based on a random forest classifier for extracting polycystic kidneys from CT volumes. Our method has been used on ADPKD patients with severe renal insufficiency, utilizing a large dataset of 55 contrast enhanced CT acquisitions.

2. RELATED WORK

In ADPKD clinical studies, commonly employed image based methods for TKV quantification are stereology [5] and manual segmentation. Both methods tend to be time consuming, especially in case of high resolution acquisitions. Recently, a new method for estimating TKV for ADPKD kidneys has been proposed [6], which does not require whole kidney segmentation, but only requires the segmentation of kidney section in its corresponding mid-slice. Another method by Irazabal *et al.* [7] describes an ellipsoid formula for TKV computation in ADPKD. Even though these methods allow fast quantification, they are likely to provide only

Part of this work was funded by TranCYST Marie Curie Initial Training Networks project, within the 7th European Community Framework Programme (FP7-PEOPLE-MCA-ITN-317246).

a rough TKV estimation which could be useful for ADPKD clinics, but not sufficient for clinical trials that need more accurate measurements.

Besides these approaches, several semi-automatic kidney segmentation techniques have been proposed previously. Fei *et al.* [8] used a deformable model based on a spline parameterization for the segmentation of polycystic kidneys from transgenic mice, but the data sets used exhibited less severe morphological changes of the kidneys. Daum *et al.* [9] proposed a semi-automatic approach based on 3D random walks for the segmentation of polycystic kidneys on T2 weighted fat saturated MR acquisitions. This approach requires manual initialization by the user for determining connected non-labeled regions based on the probability of association to the user labeled pixels. Racimora *et al.* [4] investigated a semi-automatic approach using active contours and morphological operations. Mignani *et al.* [10] used region growing segmentation with additional morphological operations and curvature based motion for assessment of TKV in T2 weighted fat saturated MRI images. An extension of their work was presented by Turco *et al.* [12]. Another recent approach by Kline *et al.* [11], describes automatic segmentation of follow-up T1 MR images from baseline acquisition. Their method necessarily requires a baseline segmentation initialization to segment follow-up MRI scans of ADPKD patients. It should be noted, however, that all these approaches focus on the segmentation of kidneys from MRI data and we are not aware of any semi-automatic method for segmenting ADPKD kidneys from CT data. In this paper, we propose a semi-automatic segmentation method based on a random forest classifier. Random forests, or more generally decision forests, have gained considerable interest in medical image segmentation. As an example, regression forests have been used for the automatic detection and localization of multiple organs in MRI by Pauly *et al.* [13] or CT by Criminisi *et al.* [14]. Focusing on kidney segmentation, Cuingnet *et al.* [15] used an approach based on regression forests for segmenting normal kidneys. Kotschieder *et al.* [16] performed semantic image segmentation using geodesic distances as an additional criterion for the node splitting in order to ensure spatial compactness of the pixel clusters of each child node. Peter *et al.* [17] presented scale adaptive segmentation approach on MR, 3D US and histological data using haar-like features sampled sequentially on a Random Forest classifier. There are many more approaches based on random forests but an extensive review of all of them goes beyond the scope of this paper.

3. METHODOLOGY

As depicted in Fig. 1(a), segmenting ADPKD kidneys is a very challenging task. In this work we propose a semi-automatic approach for segmenting ADPKD kidneys from CT data. Inspired by a recent method for estimating ADPKD volume [6], the user outlines each of the kidneys in its corre-

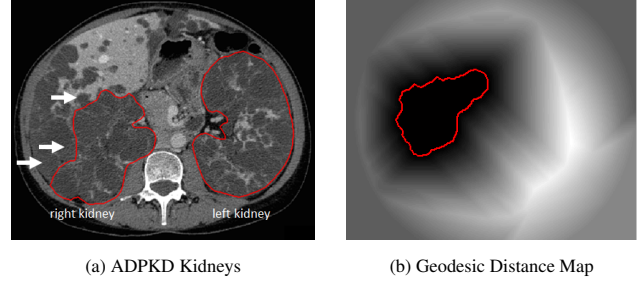


Fig. 1. ADPKD Kidney Segmentation: Panel (a) shows ADPKD kidneys. Severe morphological changes lead to difficulty in segmentation as compared to healthy kidneys. White arrows show adjacent liver cysts exhibiting similar intensity. Panel (b) shows manually segmented mid-slice along with the proposed geodesic distance map, cf. Sec. 3 for further explanation.

sponding mid-slice, i.e., the middle slice out of all the sections containing kidney. Then, we compute the intensity weighted geodesic distance to the respective mid-slice segmentation at all non-segmented voxels, cf. Fig. 1(b), according to the method of Soille [18]. This yields two 3D distance volumes, one for each kidney as additional modalities (information channels). We demonstrate that the introduction of these novel distance volumes helps to improve the segmentation results, in contrast to a random forest classifier solely trained on the CT intensities. Our goal is to formulate the segmentation task as a voxel-wise classification problem, where we assign to each voxel p a label $\ell(p) \in \{\ell_b, \ell_r, \ell_l\}$. While ℓ_b models the background class, ℓ_r denotes the label for the right kidney and ℓ_l denotes the label for the left kidney. Based on a set of labeled examples, we aim at training a decision rule by means of a random forest classifier.

3.1. Random Forests

A random forest consists of a collection of decorrelated binary decision trees. A decision tree is in turn a hierarchically ordered set of nodes, where each node has exactly 0 or 2 children – in the former case, the node is called a leaf. Using a set of labeled examples, these decision trees are trained in order to infer the relationship between visual features and labels. Concisely put, a random forest classifier provides a piecewise approximation of each class posterior over the feature space.

3.2. Feature Selection

We use so-called *box features* for classification. As depicted in Fig. 2, a box feature at a location p is defined by two offset vectors $\vec{d}_a \in \mathbb{R}^3$ and $\vec{d}_b \in \mathbb{R}^3$ which specify the centers of mass of two boxes a and b . These boxes are of size $x_a \times y_a \times z_a$ and $x_b \times y_b \times z_b$, respectively. In each of these boxes we calculate the mean value w.r.t. one intensity (information)

channel.

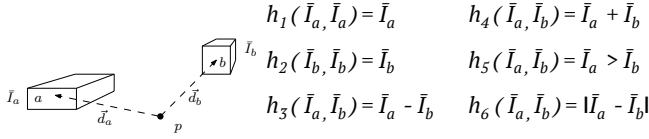


Fig. 2. Box Feature: A box feature is defined by the two offset vectors $\vec{d}_a, \vec{d}_b \in \mathbb{R}^3$, the box sizes, and the choice of the function h_i used for computing a single scalar value out of the mean values \bar{I}_a and \bar{I}_b . Note that these mean values could be computed from different information channels, i.e., the CT volume and the two geodesic distance volumes.

In particular, this means that it is possible that the mean values \bar{I}_a and \bar{I}_b can be computed from different channels to be able to capture inter-modality correlations. Once these mean values are computed, we select one of the six functions h_j , $j = 1, \dots, 6$, cf. Fig. 2, for computing a scalar feature value. Thus, one box feature can be parametrized by a vector

$$(\vec{d}_a, \vec{d}_b, x_a, y_a, z_a, x_b, y_b, z_b, k_a, k_b, j), \quad (1)$$

where k_a and k_b specify the intensity channels used for computing the respective mean value and $j = 1, \dots, 6$ specifies the function used for computing the scalar feature value.

3.3. Forest Training

Each decision tree is trained as follows: We randomly select a set of training samples S , i.e., voxels with known labels and start building the tree at the root node. At each node we select a splitting function defined by the selected box feature and a threshold value according to the following process: We randomly select 100 features and compute the minimum and maximum values of the respective feature on all samples of the node. Then we divide the range of each feature using 10 thresholds (equally spaced between the respective maximum and minimum value) and evaluate the information gain of all considered feature-threshold combinations. The splitting function at the current node is then chosen as the combination of feature and corresponding threshold which yields the overall highest information gain. Although this strategy is a greedy one, it is still one of the most popular choices due to its computational efficiency [19]. This node splitting process is repeated recursively until either the maximum depth is reached, or if the number of samples sent to child nodes is too low. Eventually, each leaf node models the class posterior estimate using a histogram from the samples that reached this leaf node.

3.4. Forest Testing

At testing (prediction) time, our goal is to classify newly observed voxels p . Thus, we start at the root and drive each

test sample through the tree according to the splits recorded at each node during the training. Finally, we obtain the class histogram stored at the leaf reached by the test sample in this particular decision tree. The overall prediction is then computed as the average of the output posteriors of each tree, and the prediction for each voxel p is then given by the class with highest average posterior.

4. EVALUATION

For our experiments, we used a total of 55 CT acquisitions from 41 ADPKD subjects. For image acquisition, a 64-slice CT scanner (LightSpeed VCT; GE Healthcare; Milwaukee, WI) was used, with single breath-hold scans and same scanning parameters for all patients (voltage 120 kV, current 150 - 500 mAs, collimation 2.5 mm, matrix 512x512, slice pitch 0.984 and increment 2.5 mm). All of these data sets were manually segmented by clinical experts in order to obtain ground truth annotations. We performed a 5-fold cross-validation, i.e., we selected 44 samples for the training and tested the trained classifiers on the remaining 11 samples. This process was repeated five times such that every data set has been used once for validation. Moreover, we performed two sets of experiments. In the first set we trained the random forest classifiers without the geodesic distance volumes as additional information channels. In the second round of experiments we trained the classifier with the same settings, but with the geodesic distance volumes as additional information channels.

5. RESULTS

We computed the average dice scores for the predicted volumes for all 55 cases. While the green bars in Fig. 3 depict the results computed with the geodesic distance volumes, the red bars show the results for the baseline approach, i.e., the random forest classifier trained solely using the CT volumes. In general, we can make the following observations: Firstly, the segmentation results tend to be better for the left kidneys which is due to the fact that the boundary between the right kidney and the liver is often hard to discriminate, cf. Fig. 1. Secondly, there are a considerable number of cases where the geodesic distance volumes improve the segmentation results - especially in case of the left kidneys. We improved in 35 out of 55 cases by 13.25% on average for the right kidney and in 36 out of 55 cases by 10.35% for the left kidney, respectively, while the baseline is only better by 2% for right kidney and by 2.42% for the left kidney on average for the remaining 20 and 19 cases, respectively. Therefore, the average gain of the proposed method (in the cases where it outperforms the baseline) tends to be higher than the average gain of the baseline method (in the other cases).

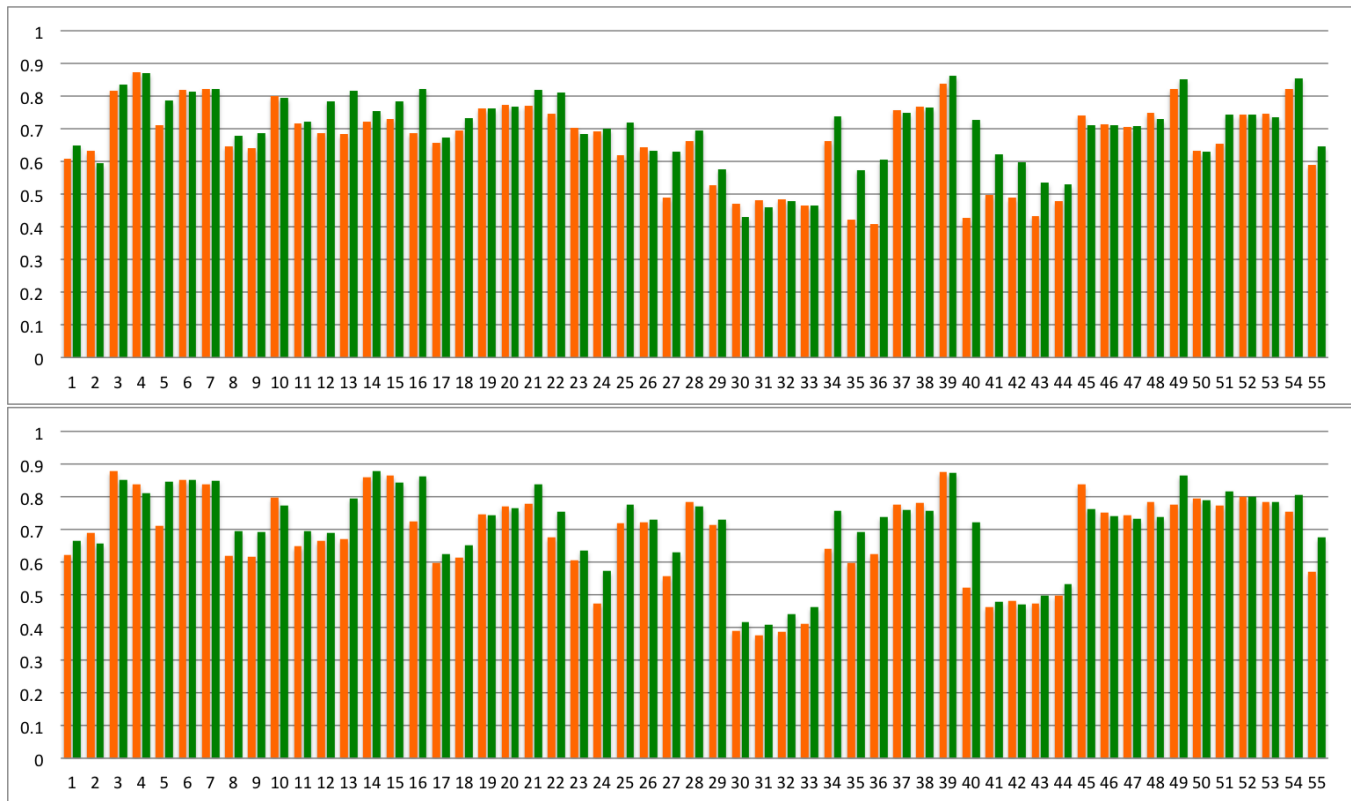


Fig. 3. Dice Scores for right kidneys (upper panel) and left kidneys (lower panel): The results obtained for all 55 acquisitions with the baseline methods are shown in red, while the results of the proposed approach are shown in green.

6. DISCUSSION AND CONCLUSION

We presented a method for segmentation of ADPKD kidneys on contrast enhanced CT. Based on the results, we may draw the following two conclusions: Firstly, the proposed usage of the random forest approach helps to improve the overall segmentation process compared to baseline approach. Secondly, all results clearly show that segmentation of ADPKD kidneys is not at all a solved task.

The main reasons are: (i) the progressive cyst expansion in ADPKD leading to a significant and unpredictable deformation and enlargement of the kidneys, especially in patients at late stage of the disease, and (ii) the aforementioned tissue inhomogeneities of surrounding organs, for instance due to neighboring liver cysts, make fully automated kidney segmentation a challenging task. As evaluating forests is computationally very efficient, it currently seems to be a good strategy to evaluate random forests. We would like to emphasize that this is to the best of our knowledge, one of the first approaches for minimally interactive segmentation of ADPKD kidneys from CT data. Both CT and MRI have been investigated for monitoring structural changes in ADPKD and for association between TKV and renal function or renal function decline [3]. The reason to acquire contrast enhanced CT

images in the current study was to perform further renal compartment measurements. But, these additional measurements are out of the scope of this paper. Even at this stage, the presented method for segmentation of polycystic kidneys can potentially reduce time and effort for volumetric measurements in clinical studies. Moreover, we hope that this work paves way for further research regarding this challenging task. Future work might include the adaption of this approach to other modalities, such as MRI, or its application to other tasks, such as polycystic liver segmentation for instance and possible improvements in the current method for a more automated segmentation approach in ADPKD.

7. REFERENCES

- [1] J. J. Grantham, “The etiology, pathogenesis, and treatment of autosomal dominant polycystic kidney disease: recent advances,” *American journal of kidney diseases*, vol. 28, no. 6, pp. 788–803, 1996.
- [2] V. E. Torres, P. C. Harris, and Y. Pirson, “Autosomal dominant polycystic kidney disease,” *The Lancet*, vol. 369, no. 9569, pp. 1287–1301, 2007.
- [3] A. B. Chapman, J. E. Bost, V. E. Torres, L. Guay-

- Woodford, K. T. Bae, D. Landsittel, J. Li, B. F. King, D. Martin, L. H. Wetzel, et al., "Kidney volume and functional outcomes in autosomal dominant polycystic kidney disease," *Clinical Journal of the American Society of Nephrology*, vol. 7, no. 3, pp. 479–486, 2012.
- [4] D. Racimora, P.-H. Vivier, H. Chandarana, and H. Rusinek, "Segmentation of polycystic kidneys from mr images," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2010, pp. 76241W–76241W.
- [5] K. T. Bae, P. K. Commean, and J. Lee, "Volumetric measurement of renal cysts and parenchyma using mri: phantoms and patients with polycystic kidney disease," *Journal of computer assisted tomography*, vol. 24, no. 4, pp. 614–619, 2000.
- [6] K. T. Bae, C. Tao, J. Wang, D. Kaya, Z. Wu, J. T. Bae, A. B. Chapman, V. E. Torres, J. J. Grantham, M. Mrug, et al., "Novel approach to estimate kidney and cyst volumes using mid-slice magnetic resonance images in polycystic kidney disease," *American journal of nephrology*, vol. 38, no. 4, pp. 333–341, 2013.
- [7] M. V. Irazabal, L. J. Rangel, E. J. Bergstralh, S. L. Osborn, A. J. Harmon, J. L. Sundsback, K. T. Bae, A. B. Chapman, J. J. Grantham, M. Mrug, et al., "Imaging classification of autosomal dominant polycystic kidney disease: a simple model for selecting patients for clinical trials," 2014.
- [8] B. Fei, C. Flask, H. Wang, A. Pi, D. L. Wilson, J. Shillingford, N. Murcia, T. Weimbs, and J. L. Duerk, "Image segmentation, registration and visualization of serial mr images for therapeutic assessment of polycystic kidney disease in transgenic mice," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the IEEE*, 2006, pp. 467–469.
- [9] V. Daum, H. Helbig, R. Janka, K.-U. Eckardt, and R. Zeltner, "Quantitative Measurement of Kidney and Cyst Sizes in Patients with Autosomal Dominant Polycystic Kidney Disease(ADPKD)," in *3rd Russian-Bavarian Conference on Biomedical Engineering*, J. Horneegger, E. W. Mayr, S. Schookin, H. Feußner, N. Navab, Y. V. Gulyaev, K. Höller, and V. Ganzha, Eds., Erlangen, 2007, vol. 1, pp. 111–115.
- [10] R. Mignani, C. Corsi, M. De Marco, E. G. Caiani, G. Santucci, E. Cavagna, S. Severi, and L. Cagnoli, "Assessment of kidney volume in polycystic kidney disease using magnetic resonance imaging without contrast medium," *American journal of nephrology*, vol. 33, no. 2, pp. 176–184, 2011.
- [11] T. L. Kline, P. Korfiatis, M. E. Edwards, J. D. Warner, M. V. Irazabal, B. F. King, V. E. Torres, and B. J. Erickson, "Automatic total kidney volume measurement on follow-up magnetic resonance images to facilitate monitoring of autosomal dominant polycystic kidney disease progression," *Nephrology Dialysis Transplantation*, p. gfv314, 2015.
- [12] D. Turco, S. Severi, R. Mignani, V. Aiello, R. Magistroni, and C. Corsi, "Reliability of total renal volume computation in polycystic kidney disease from magnetic resonance imaging," *Academic radiology*, 2015.
- [13] O. Pauly, B. Glocker, A. Criminisi, D. Mateus, A. M. Möller, S. Nekolla, and N. Navab, "Fast multiple organ detection and localization in whole-body mr dixon sequences," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, pp. 239–247. Springer, 2011.
- [14] A. Criminisi, D. Robertson, E. Konukoglu, J. Shotton, S. Pathak, S. White, and K. Siddiqui, "Regression forests for efficient anatomy detection and localization in computed tomography scans," *Medical image analysis*, vol. 17, no. 8, pp. 1293–1303, 2013.
- [15] R. Cuingnet, R. Prevost, D. Lesage, L. D. Cohen, B. Mory, and R. Ardon, "Automatic detection and segmentation of kidneys in 3d ct images using random forests," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pp. 66–74. Springer, 2012.
- [16] P. Kotschieder, P. Kohli, J. Shotton, and A. Criminisi, "Geof: Geodesic forests for learning coupled predictors," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013, pp. 65–72.
- [17] L. Peter, O. Pauly, P. Chatelain, D. Mateus, and N. Navab, "Scale-adaptive forest training via an efficient feature sampling scheme," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, pp. 637–644. Springer, 2015.
- [18] P. Soille, "Generalized geodesy via geodesic time," *Pattern Recognition Letters*, vol. 15, no. 12, pp. 1235–1240, 1994.
- [19] A. Criminisi and J. Shotton, *Decision forests for computer vision and medical image analysis*, Springer Science & Business Media, 2013.